USE OF MULTIPLE REGRESSION ANALYSIS TO SUMMARIZE AND INTERPRET LINEAR PROGRAMMING SHADOW PRICES IN AN ECONOMIC PLANNING MODEL

Daniel G. Williams

U.S. Department of Agriculture
Economics, Statistics, and Cooperatives Service

Technical Bulletin No. 1622

Use of Multiple Regression Analysis to Summarize and Interpret Linear Programming Shadow Prices in an Economic Planning Model. By Daniel G. Williams, Economic Development Division; Economics, Statistics, and Cooperatives Service, U.S. Department of Agriculture. Technical Bulletin No. 1622.

ABSTRACT

A simple method is presented for evaluating the benefit to a region (regional objective function) of new manufacturing firms. These firms are subsets of the more aggregated 4-digit SIC manufacturing industries, some included and some not, in a rural multicounty economic planning model. The model must contain many types of industries to include the full range of industry; such addition is costly. Multiple regression analysis can summarize and interpret shadow prices of export industries so that local planners in their decisionmaking can use the underlying economic characteristics of the model industries rather than use only their industry product classifications.

Keywords: Linear programming, shadow prices, objective function, multiple regression, regional exports, rural development.

Washington, D.C. 20250 May 1980

CONTENTS

SUMMARY

The benefit to a region (in terms of the increased value of a regional maximization objective function) of new manufacturing firms (or plants) can be evaluated by a simple method shown in this report: multiple regression analysis.

The firms are subsets of the more aggregated 4-digit Standard Industrial Classification (SIC) manufacturing industries, some of which are included and some of which are not included, in a rural multicounty, linear programming, economic planning model (RDAAP). The model must contain many types of industries to include the full range of local production possibilities. However, adding more industries to the model is time-consuming and costly. And, even if all 450 or so 4-digit SIC manufacturing industries were included, the results would have to be interpreted at a more disaggregated level for the thousands of firms.

The linear programming shadow prices on the manufacturing export constraints in the model reveal to local area planners the relative benefits of these various export industries. Multiple regression can summarize and interpret these shadow prices as a further aid in regional decisionmaking toward selecting the best industries for the area. Local planners can then use the economic characteristics of the model industries rather than merely their industry product classifications (SIC's). The output is also reduced in volume because the focus shifts from product type, the number of SIC's (56), and shadow prices to only nine industry characteristics. Interpreting the results in terms of these industry economic factors makes the regression analysis more useful to regional planners than programming output alone.

With certain qualifying limitations, analysts can infer the value of the generalized shadow price for an export industry or firm not already included among the given model industries. Or the regression method can be used as a summary of results only for those included industries. The multiple regression technique permits the planner to estimate directly the regression equation(s) for the specific region and model, and to insert the values for the independent variables (industry economic characteristics) for an industry "candidate" not already in the RDAAP Model. Fifty-six 4-digit SIC industries are now included. The dependent variable is the set of export industry generalized shadow prices for a given regional objective function.

Selected results obtained include: (1) industries with low transportation costs most improve the level achieved for the majority of area objectives; (2) a low industry managerial labor percentage (or whichever is the scarce labor skill factor) is generally the second largest improvement factor; (3) industries with increased value added/labor, higher value added/output, higher ratio of imported inputs to output, or lower capital/output all generally lead to improved levels for the majority of area objectives; (4) profit-type area objectives (for example, maximization of aggregate regional profits or industry regional profit rate-of-return index) often yield opposite partial-coefficient regression results and, therefore, opposite planning prescriptions from the nonprofit area objectives; and (5) for those alternative objectives which may require opposite industry economic types, explicit tradeoffs among the area objectives would have to be considered.

The results presented may apply fairly widely in many rural multicounty areas, but the users should be cautious in applying them beyond the specific

application in this report. Results also should be considered exploratory without some corroborative evidence from other sources. And, as in all multiple regression analyses, any conclusions derived from interpreting a partial regression coefficient should be considered precisely valid only with the levels for all other independent variables remaining unchanged.

USE OF MULTIPLE REGRESSION ANALYSIS TO SUMMARIZE AND INTERPRET LINEAR PROGRAMMING SHADOW PRICES IN AN ECONOMIC PLANNING MODEL

By Daniel G. Williams, Regional Economist,
Economic Development Division

INTRODUCTION

Planners of economic development in multicounty rural areas want to know what type of industrial development would be best for the planning area. A linear programming model, an optimization type of model, is often used in such planning. Linear programming is useful for solving many problems involving choice.

What specific information does the planner require for optimal area development? He or she wishes to know: (1) which industries are optimal for the area; (2) what products are produced by the optimal industries; (3) what local resources those industries require; and (4) what opportunity costs are implied by an area planning strategy suggested by (1) through (3). The researcher who understands the extensive programming printouts detailing optimal activity levels and shadow prices learns a great deal from linear programming studies. But it is often difficult to explain the results to a layperson, area official, or planner either because there is too much detail to absorb or because these people use other languages and concepts for the same economic or social problem.

In the past, this translation from the language of abstruse computer printouts—a form perhaps more easily understood by the research specialist—to that of user-oriented information, has depended upon the patience, astuteness, and ability of the researcher. This bulletin illustrates how multiple regression analysis can be used by technical staff in the planning office to help summarize, translate, and reduce in volume the linear programming results into a user-oriented framework. This interpretation is extended into a statistical or inferential context. The multiple regression procedure assumes that the initial economic or social problem was correctly solved by the linear programming algorithm, and it involves an examination of how regression analysis of these programming results can help the researcher or planner to summarize, convert, and interpret them.

The planning task will be easier if not only the type of industry "selected" by the model's optimal solution is known (that is, the SIC code number), but also the economic characteristics of optimal industries. Is the industry light or heavy, capital or labor intensive, polluting or nonpolluting, high wage or low wage, and so on?[1]

[1] Obviously, not all these questions are answered by the nine industry economic characteristics used in this analysis. However, there is no conceptual reason why they could not be so answered. The proper input-output coefficients would need to be included in the linear programming model, and the respective industry characteristics would be calculated from these coefficients.

As a hypothetical example, the optimal solution to the linear programming model might indicate that the manufacturing industries of meatpacking, poultry processing, refrigeration machinery, and so on, should be located in the region to maximize the increase in gross regional product. For the planner, multiple regression analysis would identify and estimate the size of economic characteristics of optimal industries (and of industries in the model which are *not* in the optimal solution) which, other things being equal, lead to greater or lesser increases in gross regional product. For example, low transportation costs or high value added per unit of output might be industry characteristics associated with larger increases.

The multiple regression method is illustrated with an activity analysis, economic development model (RDAAP) used for economic planning in multi-county rural areas.[2] The forerunner or precursor to this current model is the Kentucky Model, developed by Robert G. Spiegelman, and others (*4*).[3]

The nine industry economic characteristics into which the programming results are transformed depend upon the type of activities in the model. In general, such characteristics should consist of only those factors which are easily understood and can be calculated readily from the model input. Also, such factors should be presumed to have, relative to other possible factors, a more significant linear functional relationship with the linear programming output.

The results reported here can be viewed in three senses:

- They describe and summarize only the 4-digit SIC industries included in the model.
- They can be applied (inferred) to nonincluded[4] model industries.
- They can be applied (inferred) to actual (or potential) area industries.

---

[2] The model referred to is the Rural Development, Activity Analysis Planning Model (RDAAP). It is applied to a three-county area in northwest Arkansas—the BMW Region (Benton, Madison, and Washington counties)—using 1960 as the base year of the plan and 1970 as the target year. One of the six versions of this model, the RDAAP Model-Basic Model, is designed to ascertain what the planning implications *would have been* for 1970 so that the results can be compared with the *actual* results that year. The examples in this bulletin, however, are taken from another model version, the RDAAP Model-Adjusted Planning Model, which was constructed to improve the model in a planning sense, from the Basic Model. The RDAAP Model involves a 10-year time frame, but is concerned specifically with only the terminal, or 10th year of that frame. Constant 1963 dollars are assumed.

Although Benton and Washington counties together were declared an SMSA after the 1970 U.S. Census of Population, and therefore can no longer be considered rural, this application of a rural model can provide a glimpse into how an area should evolve (optimally) from a relatively more rural to more urban status.

[3] Three papers by the author explore more deeply other interesting and specialized areas of this current research (Williams *5, 6,* and *7*). Two other manuscripts describe certain aspects of research on the entire model. The first (Williams *8*) will summarize some of the more important research results, while the second explores the components, framework, and mathematical structure of the model (Williams *9*). Italicized numbers in parentheses refer to items in References at the end of this report.

[4] That is, those industries not among the 56 4-digit SIC manufacturing industries included in the linear programming model.

There is a fine line of distinction between the second and third senses. The second sense refers to use of secondary data for model industries, which may or may not be a random sample from the planning area. If it is not, inference can still be made "within" the model to other industries not in the model, but whose industry input-output coefficients are also estimated from secondary data. Similarly, inference can be extended to industries included in the model but which become nonincluded when one or more industry coefficients is changed, leading to a change in the level of an industry economic characteristic(s). However, inference (second sense) in either of these examples does not necessarily imply validity for the actual planning area. Accordingly, the third sense refers to whether the "model" results are valid for the actual planning area (the "real world"). These three senses are in order of perhaps a decreasing statistical validity, but an increasing area planning utility. Nie (*3*, p. 321) describes the two main interpretations of multiple regression analysis: Summary (descriptive) and inferential. In this report, which emphasizes rural, small-area economic development planning, the predictive or inferential senses are stressed.

In both of the two inferential senses given here, all firms or plants, as subsets of the 4-digit SIC classifications, can be considered. Because these plants generally will differ in input-output coefficients from their 4-digit SIC "average," the linear programming model's selection of an optimal 4-digit SIC industry may or may not mean that a particular plant or firm in that same SIC is also optimal.

The linear programming model's solution consists of output and usage levels for industries, labor skills, labor commuting, and shipments (exports) outside the BMW Region. The model includes activities which produce manufactured products that can be used for local consumption or for export from the region. Constraints limit the amount of export that can be delivered (1) to a nearby market at one set of transportation costs (inner ring), and (2) to a more distant market at a higher set of transportation costs (outer ring). The question here is: Which export activities would planners prefer to attract to their region? These exports consist of 101 separate export activities in the RDAAP Model, which represent 56 different product types corresponding to 56 4-digit SIC manufacturing industries. This number (56) expands to 101 because exports to the outer ring are differentiated from those to the inner ring, and 45 of the 56 industries are considered to have export potential to both areas.[5]

The multiple regression analysis reduces and translates the linear programming results from 101 export product types to nine (or fewer) industry economic characteristics, such as capital/labor, value added/labor, and transport charge,[6] which are perhaps in a form more intelligible and useful to an area

---

[5] That is, exports sent to the inner and outer export rings are considered different commodities for the same type of industry. Moreover, in the RDAAP Model, some manufacturing industries do not export at all, and some export to the inner ring but not the outer ring.

[6] Note that for the 45 outer ring export industries, only the transport charge, and none of the other independent variables, will differ in value among the nine characteristics when one compares an industry shipping to the outer ring with its counterpart SIC industry shipping to the inner ring.

There are no high pairwise correlation coefficients between independent variables (industry characteristics), nor any obvious high multivariate correlations, both necessary requirements to lessen the possibility of multicollinearity. There is also not likely any problem with auto or serial correlation in the error terms because of the high value observed for the Durbin-Watson statistic.

planner than would be details by SIC code. The volume of data is smaller and more manageable; only nine (or fewer) "bits" of information will result rather than 101. Instead of viewing the regression scheme, which uses the linear programming results, as providing a *substitute* for those results, one can view it as an adjunct to the results, in increasing understanding and providing a useful summary. The industry economic characteristics would probably reflect more the language of the practitioner (such as the regional planner) and less the language of the researcher. In other words, this technique presents a way to convert more abstruse research results into ideas more easily understood and useful to planners, councils of government, chambers of commerce, elected officials, or other area functionaries.

The cost, or information lost by this data reduction (and translation) scheme, can be understood by recognizing that the linear programming model explains 100 percent[7] of the output[8] variability; that it implicitly represents a multiple R^2 of 100 percent. The nine (or fewer) explanatory variables (industry economic characteristics) explain a smaller R^2 portion of the variation in generalized shadow price by export (defined in next section). The reduction in multiple R^2 from 100 percent represents this cost. In other words, reducing the precision of knowledge of the generalized shadow price (for example, through this multiple regression technique) enables a reduction in the number of variables to be considered (9 industry characteristics) and a translation of the original variables (101 product types) to another form (SIC's to industry economic characteristics). The size of the multiple R^2 measures how good[9] this reduction and translation scheme is relative to the complete linear programming results.

OBJECTIVE FUNCTION AND THE GENERALIZED SHADOW PRICE

Consider the objective of maximizing gross regional product. The manufacturing industry[10] export constraints (rows) yield valuations which are labeled shadow prices. A shadow price for a particular export row measures how much an objective function value will improve for a one-unit ($1 million) increase in

[7] For example, with the assumption that all coefficient values for the linear programming model are fixed, the linear programming algorithm yields deterministically and exactly the valuations on the resource constraints—the shadow price values.

[8] In particular, the generalized shadow prices of the manufacturing export row constraints. The concepts of shadow prices and generalized shadow prices will be explained in the next section.

[9] In terms of the accuracy of the multiple regression equation's estimate of the generalized shadow price versus the linear programming model's "estimate" (that is, 100-percent "accuracy").

[10] Service industries are assumed to have no export possibilities from the BMW Region. Raw (nonprocessed) agricultural produce can be exported directly, or processed locally if the corresponding food processing (manufacturing) industry is included among model production industries. Only this type of agricultural export is included among the exports considered in this analysis.

the constraint level for that export. Thus, manufacturing export industries can be ranked according to their relative desirability by use of their shadow prices. Nonzero shadow price levels will result for industries which export at the (upper) constraint limit. Those which export below this limit (but not at a zero level) will have zero shadow prices, as will industries which exhibit zero export levels in the optimal solution. For the latter, the solution yields a "reverse" or "negative" shadow price called the reduced cost. This price measures how much the objective function would be worsened if these nonoptimal export industries (individually) were to enter the solution at a unit export level.

In short, *all* manufacturing export industries, whether or not they enter the optimal solution at a positive level (produce or not produce for export), can be ranked according to their shadow price desirability. The "true" shadow price export rows (i.e., corresponding to positive levels of exports) in terms of decreasing relative worth—can be ranked from high absolute value down to zero. The zero level exports can be ranked in terms of their reduced costs from low absolute value to high, also in terms of decreasing relative worth to the solution.[11] The true shadow prices, together with the reduced costs, are labeled here "generalized" shadow prices. These prices represent the dependent variable in the multiple regression analysis. For simplicity, unless otherwise stated, shadow price will refer to generalized shadow price.

THE MULTIPLE REGRESSION AND LINEAR PROGRAMMING MODELS[12]

Because the linear programming model (RDAAP) is of the usual form for such models, its mathematical structure will not be presented.[13] Several aspects of its use in multiple regression analysis should be noted, however. First, the industry (input-output) coefficients can be viewed as though they are *fixed* in the linear programming analysis (even if they are not). Thus, the nine industry economic characteristics calculated from them similarly can be considered to be fixed. These characteristics (table 1) are the independent variables in the regression analysis, but do not have to be considered as random variables. Interpreting the independent variables as being fixed or pre-specified is valid in regression analysis even with the assumption of randomness in the independent variables.[14]

[11] For example, for a maximization objective function, the true shadow prices run from high positive values down to zero, while the reduced costs are listed from low negative values to high. Both are ranked in terms of decreased worth.

[12] For a more thorough discussion of the *theoretical* underpinnings of this research involving a linear programming model in connection with a multiple regression analysis, see (*10*).

[13] The RDAAP Model includes service, manufacturing, and government sectors, and an agricultural sector in which technological progress is simulated by conversion of regressive farms into progressive farms. See (*9, 6*).

[14] See (*2*, pp. 106-138). It is not exactly true to say that the independent variables in this analysis are pre-specified as the industry coefficients in the linear programming model used were "given," and, thus, the values for the independent variables also were given. Nonetheless, the ranges of the independent variables seem sufficient for an acceptable experimental design.

Table 1—RDAAP Model's manufacturing export industry economic characteristics: Independent variables

Independent variable	Unit	Calculation
Transport costs per dollar of export	$\$10^6$	Transport costs per million dollars of export
Capital/output ratio	–	10-year capital ÷ 10th-year output
Capital/labor ratio	\$100/hours worked	10-year capital ÷ 10th-year labor ("current" account)[1]
Rate-of-return ratio	–	10th-year total profits ÷ 10-year capital
Value added/output ratio	–	10th-year value added ("total" account)[2] ÷ 10th-year output
Value added/labor ratio	\$100/hours worked	10th-year value added ("total" account) ÷ 10th-year labor ("total" account)
(Managerial labor)/(total labor) ratio	–	Managerial labor ("total" account) ÷ total labor ("total" account)
(Skilled labor)/(total labor) ratio	–	Skilled labor ("total" account) ÷ total labor ("total" account)
Imported input costs per dollar of output	$\$10^6$	Direct imported input costs ("total" account) per million dollars of output ("current" account)

[1] For "current" goods production, rather than "capital" goods production.
[2] For "current" vector, plus associated "capital" vector (assumed to be 15 percent of total 10-year capital requirement) for the 10th year.
– Dashes mean that the associated variable has no units. It is a pure ratio.

One must assume the *conceptual* possibility of repeated sampling, however, with the pre-specified variables, for a meaningful interpretation of the random error term and the statistical F tests for both the whole equation (multiple R^2) and also the individual partial-regression coefficients, B (and BETA), the latter measured in standard deviation units. The hypothetical sampling can be visualized as being repeated for many alternative data input sets for the given linear programming model, the only requirement being that the 101 export types (SIC's) remain in the model. When one uses common distribution assumptions (*2*, pp. 25-29, 133), the nine independent characteristics can be considered to have remained fixed in value; only the generalized shadow prices (dependent variable) and the error term vary in each case.[15] The only industry input-output

[15] This, of course, is only a "hypothetical" or conceptual sampling scheme, not actually implemented, although necessary for the concept of a random error term. That is, once the regression equation has been estimated with the fixed independent variables, inference can be made on other industries or firms with different, although again fixed (that is, not random) economic characteristics.

coefficients "permitted" to change between samples would be those which do not affect the levels of the nine characteristics. One then assumes the resultant changes in shadow prices are due to the changes in the (hypothetical) 10th and other characteristics not considered in the regression model, but which, in general, could change with the sampled industry input-output coefficient changes. The error term "represents" these 10th plus characteristics and, by the Central Limit Theorem, it is assumed to be a random variable with zero mean and constant variance.

The least complex method to compare generalized shadow prices with industry economic characteristics might seem to be to rank model industries by the nine industry characteristics. Some obvious correlation may occur between the generalized shadow price rank and the industry rank by industry characteristic. However, except for the few industries ranked top and bottom as to generalized shadow price, this *ad hoc* technique proves unwieldy and unworkable. It would be extremely difficult, if not impossible, to discover the multivariate effects in this way. Also, because a number of objective functions are examined, the model can yield a different generalization corresponding to each objective function as to what is best from the planner's point of view. Therefore, this *ad hoc* ranking procedure can yield different results dependent on the particular objective function chosen. Although multiple regression analysis, too, would yield different results, the simplicity of the technique enables many alternative objectives to be handled almost as easily as one.

Such an *ad hoc* procedure yields a "mass of detail," from which it seems difficult to discern any consistent patterns. Multiple regression analysis is one statistical technique useful in reducing the data to one underlying regression equation (assumed linear) from which these patterns can be observed more easily. Use of multiple regression seems preferable to this researcher, rather than adding large numbers of 4-digit manufacturing industries to the model. It seems advisable not to add more SIC's, but to learn characteristics of the optimal industries, and to reveal those characteristics which seem to improve the regional objective function. Given restricted time and research funds, such a course will cost less in time and money.

The independent variables considered for analysis were chosen partly based on ease of calculation, but only those variables presumed or hypothesized to influence the dependent variable were included. To lessen the possibility of multicollinearity, we reduced the number of these independent variables to 15, by eliminating any variables which, *a priori*, were thought to be highly pairwise correlated with included variables. The final number was reduced to nine variables, after initial runs still showed some high pairwise correlations between some of the independent variables.

These nine variables (table 1) measure common industry economic characteristics. While they are by no means all of those which conceivably could have been used, it is hoped that these characteristics can "explain" a fairly large fraction of the multiple R^2 of the regression equation. The general hypothesis, of course, is that economic characteristics of industries do influence the generalized shadow price. The regression results will be compared with what one might have expected from economic theory.

INTERPRETATION OF THE MULTIPLE REGRESSION RESULTS

The interpretation or use for the regression output can be visualized in the three senses referred to earlier:

- Descriptive or summary only of those industries included in the model.
- Applied to industries not included or slightly changed from those in the model.
- Applied to actual (or potential) area industries.

Statistical Inference: The First Two Senses

The statistical analysis of the regression results applies easily to the first sense, in which only the industries in the model are considered; the question is ignored of inference to nonincluded model industries (or the same SIC's as in the model, but with changed industry economic characteristics). However, there may be some question as to whether this analysis yields a statistical basis for the other two senses.

In the second sense, it is most legitimate statistically to predict a shadow price for an (SIC) industry already included in the model, but with only one or a few modest changes in the input-output coefficients from those in the sample, leading to a change in one (or a few) of the nine industry characteristics. The 10th plus characteristics would then be, in general, only slightly altered. For a nonincluded SIC model industry (or firm, plant, and so on), an inference from the sample as to its shadow price would be more accurate if it too were similar to those industries already in the model. For example, if the values for the economic characteristics of the proposed industry lay in the same range as those in the model sample, it seems likely that characteristics 10th and above would also lie in a similar range. The error term would therefore be of similar size and characteristics to those in the sample. Such an occurrence might be likely if the data for the proposed industry were drawn from at least a somewhat similar data set as that of the original sample. Thus, inference could be made as to the shadow price of the new industry.

One might criticize this procedure because there is no certainty that the results from such a regression for a given sample of export industries (56 industries in the RDAAP Model) would be valid enough to infer a shadow price result for a *nonincluded* model export industry.[16] Because of the "geometry" of the situation, one might not know, *a priori*, how a new constraint (for example, a new export industry row) would affect the feasible region of the linear programming problem and, thus, the proposed industry's shadow price.

Nonetheless, one would have some idea about the shadow price of an added industry whose industry economic characteristics are drawn from a data set similar to that of the original sample. Consider the following argument. Sup-

[16]This criticism would apply even to a plant or firm within an *optimal* 4-digit SIC industry in the model, dependent on how much the plant differed from its more aggregated industry classification. The necessity for uncovering the underlying industry characteristics would still apply, so that a proper choice could be made of a plant or firm within that optimal SIC.

pose that several SIC export industries were added to the model (for example, for a total of 57-60 industries), and that their respective shadow prices were not *a priori* estimable from their industry characteristics based on the multiple regression results obtained from the original sample of 56 industries. This would imply that regression results from the 60 industries would differ substantially from those of the 56. In other words, the implication is that the industry shadow prices and rankings would differ widely between the two sets of industries. If so, not only would the regression analysis proposed here prove invalid, but also such a criticism would seem to question the validity and usefulness of *all* of the linear programming models and shadow price planning analysis. If model results are, in general, extremely sensitive to minor changes in the number of model industries selected, only including *all* of the nearly 450 4-digit SIC manufacturing industries in the RDAAP Model would give the shadow price results any *general* validity. Even then, firms or plants within those SIC's would have to be evaluated. Yet, the validity problem would remain, as export constraints are only one of many groups of constraints in a programming model. Put more simply, the regression analysis would seem to have no more or less fundamental validity or utility than linear programming itself.

The results from the regression analysis, however, should be considered valid only at or near the optimal corner solution,[17] where opportunity costs implicit in the model do not change, or change only slightly. As will be explained shortly, the individual regression coefficients are *partial* coefficients. They apply only to small changes in the values of the independent variables and to an unchanged underlying population which the sample purports to estimate. To satisfy both of these prerequisites, only a few industries similar in range of values to the independent variables in the sample, should be added to the set of industries already in the model.

Adding too many industries to the group of model industries might lead the programming solution excessively far from its original optimal solution, and perhaps to another corner solution (the new optimum) with vastly changed opportunity costs. In short, the export shadow prices might change radically. Therefore, the shadow prices of the vastly expanded industry group could not be predicted by multiple regression on the original set of industries.

The basic reason for limiting the number of added industries is not so much the number itself, but rather the increased likelihood in such a sizable industry group of industries whose industry characteristics lie outside the bounds observed in the original sample. These industries would likely use scarce resources in a different pattern from the original model industries, and change the opportunity costs substantially. Nevertheless, if one keeps in mind these above restrictions, and considers industry candidates accordingly, this multiple regression technique should prove useful to area planners using programming models.

Of course, the programming model need not actually be rerun with the several "new" industries (those nonincluded model industries for which shadow prices were inferred to be high), although they could be added to check the validity of the regression results. About 400 4-digit SIC manufacturing industries were not examined in this analysis, and the number of firms or plants within those more aggregated industries probably numbers in the tens of thousands.

[17] According to the theory of linear programming, an optimal solution is found among the finite number of "corners" of the geometric feasible region (in n-dimensional space).

Calculating an *exact* inferred shadow price for each excluded industry would take time. However, making approximate judgments, using only some of the industry characteristics, would probably be much simpler than creating and running a linear programming model of thousands of industries, firms, and plants, and ascertaining directly the exact value of each industry shadow price from the optimal solution. Such a process to examine individual industries quickly can be viewed as an approximate decision-rule to estimate shadow prices.

The large group of excluded model industries can be pared down to a final selected small group which probably will be at least "near best" among the excluded model industries. This method seems preferable to (1) ignoring excluded model industries altogether, or (2) including all possible industries in the programming model.

To further validate the proposed regression technique, the analyst can test the model for its sensitivity to changes in the number and sizes of constraints in general, and added export activity constraints in particular. A direct test would be to add (or remove) a few industries and see if the shadow price results change and, most importantly, if the *rankings* change. If it is difficult to add industry activities commensurate with those already in the model, perhaps "artificial" vectors could be created, high in characteristics which better the objective function, and low in those which worsen it, as determined from the prior regression analysis. The new activities should yield high shadow prices. Adding these vectors to the model, and rerunning it with the new industries, should determine whether the predictions were correct. None of these tests has been attempted here because of a lack of time; they would be a useful endeavor.

A more indirect test, however, has been done: Multiple regression results from the Basic Model version of RDAAP were compared with those of the Adjusted Planning Model version. Only the Adjusted Planning Model results are discussed here. Both model versions contain the same number of export activities and constraints, and levels of such constraints; other types of constraints, constraint values, and activities differ between them.

If the programming model results were extremely sensitive to small changes in both the number and types of constraints, it probably would be revealed in substantially different regression results for the Adjusted Planning and Basic Models. In general, however, the multiple regression results were quite similar for the two model versions. The labor skill supply coefficients lead to a notable exception, however, because labor of one skill type is assumed to be in shorter supply in one model, and labor of another skill type is assumed to be in shorter supply in the other model. As would be expected, the regression results (B, BETA and F values) for the two models differ for these two labor skills, whose ratios are included among the industry economic characteristics.

Statistical Inference: The Third Sense

Inference in the third sense described refers to whether the model results provide a basis for planning in the "real world." Applying these results *outside* the model will not be as secure theoretically or statistically as predictions *within* the model. But going outside the model to the real world is what planning is about. Hence, this report, which emphasizes the regional planning use of this multiple regression technique, although focusing upon inference in both these latter two senses, will especially be concerned with the third sense.

All industries in the RDAAP Model were created from secondary data and do not reflect a random sample of firms (or potential firms) in the area.[18] The question of a statistical basis for planning (third sense) cannot, then, be answered precisely. One could have conducted such an industry sample in constructing the linear programming model, and would know the statistical confidence intervals surrounding the industry (input-output) coefficients in the sample. By extension, the statistical confidence intervals for all shadow prices and reduced costs would perhaps also then be known. If this were done, an interpretation (third sense) as to the regression tests of significance might be somewhat more valid for use in area planning. All variables, dependent and independent, would then be considered random.

However, as usual in formulation of linear programming models, a random sample of firms was not made. The question of the accuracy of these regression results (the third sense) would apply not only to these regression results but also to the programming results of *any* linear programming model that used secondary or nonrandom sampled data. In this respect, the proposed regression technique has no less validity for area planning than the results from most linear programming analyses. If the programming model input (and therefore output) is thought to represent the multicounty planning area sufficiently, the multiple regression results will do so also. Similarly, if such linear programming results are believed to be valid for another study area, the regression technique is validated too. Parametric programming on "questionable" input-output coefficients, or use of stochastic programming are two ways to approach the problem of statistical reliability of the input data, although the latter method was not used in the linear programming analysis for the RDAAP Model.

The following section, written from the viewpoint of the regional planner, implicitly focuses on inference senses two and three for area planning. It also provides further rationale for undertaking the regression analysis.

THE REGIONAL PLANNER

Using the export row generalized shadow prices discussed earlier, the regional economic planner can determine the relative (and absolute) desirability of the export industries in relation to the welfare (objective function) of the planning region. The planner can suggest to area officials or the chamber of commerce and the like, which are the better types of industries to locate within the planning region. In other words, the relative rankings of the shadow prices reveal, other things being equal, which industries should be induced to locate within the region, and which industries should not.

The question of whether any industry or group of industries actually will locate in the planning area is not directly addressed here although it is suggested that those industries which lead to higher profits or rates of return for the region would likely be relatively more attractive to entrepreneurs. However, those

[18] For example, manufacturing data are taken from "ruralized" (non-SMSA) data from the "worksheets" for the 1958 U.S. input-output table prepared by the U.S. Department of Commerce.

industries originally chosen for inclusion in the programming model should be selected on the basis of whether the planner feels they will be economically suitable for the area so that the linear programming results, and the regression results derived from them, will yield more useful planning advice. This bulletin explores which are the best industries for the area for each alternative regional goal or objective. The regional planning strategies needed to attract these "optimal" industries to the region are, of course, a necessary adjunct to the research contained in this bulletin, but they are not discussed here.

The RDAAP Model, however, includes only a limited number of export industry choices (56 of a total population of about 450 4-digit SIC manufacturing industries). How, then, can the area planner state that a given group of industries is best for the area when this planner has been given an incomplete set of choices?[19] Accordingly, the analyst or planner would wish to compare 4-digit SIC industries not included in the model or to compare firms or plants at a more disaggregated level within included or nonincluded model industries, to determine their value (generalized shadow price) to the model (sense two) or to the actual planning area (sense three).

Because the RDAAP Model includes only a few export industry choices, increasing the industry scope in the linear programming model would be desirable if there were no costly expense in creating and adding additional industries to the model.[20] But the true area planning task involves firms or plants, not 4-digit SIC's. Hence, an extremely unwieldy and costly linear programming model

---

[19] As explained in footnote 2, the RDAAP Model-Basic Model compares model results (employment increases) with the actual increases observed for the BMW Region between 1960 and 1970. About 50 percent of the 4-digit SIC manufacturing industries which grew in the region over this decade did not exist as potential local production activities in the model. While a precise SIC matchup of optimal and actual industries is not necessarily required, or even desirable, for the area—actual development may differ from what would have been "best"—at least the *possibility* of such a matchup should be included in the model.

But, as the RDAAP Model manufacturing matrix was designed to be applied to many alternative rural, multicounty areas, the probability of such a matchup is rather remote, in general, for any one area. Hence, there is a need to increase the number of manufacturing activities in the model if it is to have a more universal area application. The purpose here, however, is to show that this expansion may not be necessary because of the switch in emphasis from SIC's (product type) to industry economic characteristics. Industries could then be evaluated without adding them directly to the model.

[20] For the RDAAP Model, there was a problem in adding to the number of 4-digit SIC industries. The methodology for neither the current nor capacity expansion activities for service and manufacturing was documented in sufficient detail by R. G. Spiegelman and others of the Stanford Research Institute (SRI), in their revision of the model after the Kentucky Model. One could not replicate their procedures to learn if additional vectors created would be compatible with those already in the model. These vectors created by SRI form, in general, the core of the manufacturing and service matrix used in the RDAAP Model.

This documentation of the SRI model research after the Kentucky Model is given in five (unpublished) reports: Spiegelman, R. G., and E. W. Lungren, *Generalized Model for Rural Development Planning,* USDA Contract No. 12-17-09-1-398, Progress Reports I, II, III, and IV, June 1968, October 11, involving thousands of firms (or plants) would be needed to cover manufacturing fully. The regression results can provide at least some progress in this respect by use of the sample of 56 SIC industries (101 export activities). To use this shortcut, the analyst or planner must substitute the concept of industry economic characteristics for that of industry product type (for example, SIC).

As a hypothetical example of the use of the shortcut, a region's planners would not focus on whether to attract an industry producing axe handles rather than one producing poultry, but whether to attract industries with low transport costs, high import requirements, low capital intensity, and so on. The translation to industry characteristics reduces the mass of linear programming output and puts it in a form much easier for the regional planner to understand and use. In any case, the sample results obtained here will be at least rough indicators of possibly more general results.

MULTIPLE REGRESSION RESULTS: GENERAL REGRESSION METHOD

Specific results of the model are assessed in this and subsequent sections. These results are examined for eight alternative regional objectives, which makes the regional planning problem more complicated. With such a multiplicity of objectives, two or more area objectives of roughly equal importance to regional officials can produce conflicts. Diametrically opposite industry types (as to characteristics such as capital/output, and so on) may be recommended corresponding to each respective objective.

The first multiple regression method used is the "general" method.[21] In eight successive regressions corresponding to each of eight different regional objective functions, the nine independent variables are related to the generalized shadow prices associated with each respective area objective function. Table 2 lists the eight alternative objective functions to be maximized in the RDAAP Model-Adjusted Planning Model. In the first two maximands, consumption and Government spending are constrained to attain at least certain minimum levels. The other six maximands have no similar constraint minimums, but they preserve the requirement that aggregate tax and wage bill levels generated determine the levels of Government spending and consumption, respectively. Calculation of the first seven regional objectives is fairly straightforward. To calculate the last objective function, however, we weight each separate industry by its individual profit rate of return. We multiply each rate times the industry production level (in million dollars) in the optimal solution, and sum over all industries.

1968, February 1969, and April 1969, respectively; and Lungren, E. W., *User's Manual For Activity Analysis Model,* prepared for the U.S. Department of Agriculture, February 1968. For many aspects of the RDAAP Model's structure, however, the methodology and activities have been altered or expanded beyond those developed by SRI so that fairly substantial portions of these SRI reports are no longer valid for the present model (RDAAP).

[21] Later a "stepwise" procedure will be discussed in which only statistically significant explanatory variables remain in the results. The variables are added one at a time until all (and only) statistically significant variables have been entered.

Table 2–RDAAP Model generalized shadow prices: Dependent variables

Unit	Dependent variable (related objective function)
$\$10^6$	Regional balance-of-payments surplus
$\$10^6$	Regional balance-of-trade surplus
$\$10^6$	Gross regional product
$\$10^6$	Local[1] regional value added
$\$10^6$	Local regional aggregate wage bill
10^4 hours worked	Local employment total
$\$10^6$	Regional profits total
$\$10^6$ (index)	Regional rate-of-return index

[1] Includes no imputed values for labor commuting in or out of the Benton, Madison, Washington (BMW) Region.

The results obtained by looking successively at the *pairwise* or simple correlations between each dependent and independent variables are not listed here because of their perhaps dubious utility for planning purposes. Some of the pitfalls into which one can fall when using "simple" correlation coefficients uncritically will be mentioned. For example, the pairwise correlation coefficient might show more (or less) "strength" or size than the partial regression coefficient. Also, its sign conceivably could be reversed if other variables more than "cancel out" the partial effect.

To use the simple correlation, a planner must assume that the *other* independent variables will always behave similarly to those in the sample from which the regression is obtained, when considering some industry not in the model for its potential value to the objective function solution. Because the partial coefficient measures the effect of the independent variable with the values for the other independent variables remaining unchanged, the above assumption is not needed. Thus, the simple correlation can be used, but only with extreme caution.

For most of the eight objective functions, the amount of multiple R^2 regression explanation of the variation in generalized shadow price is large. In the least successful regression–maximizing total regional profits–the nine variables explain nearly half the variation in generalized shadow prices (table 3). Two of the adjusted R^2 figures[22] are above 79 percent; four are above 70 percent. Also, the regression results for all eight objective functions are statistically significant (F test) for the whole equation at far better than the 0.01 level.

What do these results show? Namely, that the attempt to search for and find a relationship between generalized shadow prices and industry economic characteristics has been reasonably successful. And, translating the RDAAP

[22] "Adjusted R^2 is an R^2 statistic adjusted for the number of independent variables in the equation and the number of cases. It is a more conservative estimate of the percent of variance explained, especially when the sample size is small" (*3*, p. 358).

Table 3–Multiple regression results, multiple R^2: General and stepwise regression methods[1]

Method	Dependent variable (related objective function)							
	Regional balance of payments	Regional balance of trade	Gross regional product	Local value added	Local aggregate wage bill	Local employment	Regional profits	Rate-of-return index
General regression method:								
Multiple R^2	0.83524	0.81177	0.73996	0.61458	0.68891	0.69233	0.43036	0.76668
Adjusted multiple R^2	.81894	.79315	.71424	.57646	.65814	.66190	.37403	.74361
F statistic[2]	51.25627	43.60472	28.77217	16.12293	22.39084	22.75189	7.63899	33.22517
Stepwise regression method final step:								
Multiple R^2	.83464	.80954	.73337	.60108	.68383	.69176	.40713	.76352
F statistic[3]	58.04	56.47	43.09	28.63	33.88	29.82	13.05	42.90

[1] 101 observations and 9 independent variables.
[2] Degrees of freedom: 9,91. Critical F = 2.64 for 1-percent significance.
[3] Degrees of freedom (varies left to right): 8,92; 7,93; 6,94; 5,95; 6,94; 7,93; 5,95; 7,93.

Model solution of SIC categories and shadow prices into industry economic characteristics can provide planners with both a useful summary of the model results, and more information for decisionmaking.

The results listed in tables 4-7 represent perhaps the most important aspect of the regression analysis, for they contain a list of the regression output for each individual independent variable. These are now discussed.

MULTIPLE REGRESSION RESULTS: GENERAL REGRESSION METHOD—PARTIAL COEFFICIENTS

The regression coefficients in tables 4 and 5 are sample partial coefficients which measure the change in the value of the dependent variable in the regression equation for a unit change in the particular independent variable, given that the rest of the independent variables remain unchanged. The B coefficient (table 4) measures this change in original units—units in which the objective function is measured as reflected in the units of the shadow price. The BETA coefficient (table 5) measures the same effect in standardized units. That is, it measures the change in standard deviation units (of dependent variable) per standard deviation unit change (in independent variable).

Comparisons among independent variables (and among objective functions) as to the size of the partial coefficients are easier and more meaningful when standardized coefficients are used. The partial coefficient results will be discussed only for variables which show a statistically significant F value of 25 percent or better. Many independent variables were found to be much more significiant than at the 25-percent level.

Most importantly, all results in this section, and many in the Conclusions section refer to the partial coefficients. Thus, any statements or interpretations of these results should be understood to hold strictly true only with values of the other eight independent variables remaining unchanged. The results may be more general, but this aspect was not explored here.

Table 7 shows the number of times (for eight dependent variables) that each independent variable exhibits a statistically significant[23] B (BETA) coefficient, and the number of such significant variables for each separate objective function. Also, for these variables, the BETA coefficients are relisted from table 5, and ranked as to absolute value (size) within each successive objective function column. The capital/output, value added/output, value added/labor, and (managerial labor)/(total labor) ratios are significant for all eight objective functions. Transport cost per dollar of export is significant seven times, and the imported input cost ratio is significant six times. The other three independent variables are statistically valid for only three objective functions or less, but every variable is significant for at least one objective.

What are some general observations on the statistically significant regression coefficients given in table 4-7? Table 6, listing the statistical F values for the partial coefficients, shows that transportation cost is, over the majority (six) of the objective functions,[24] the most statistically significant independent variable. Table 7's statistically significant BETA (standardized) coefficients can be used to determine the relative rankings of the significant independent vari-

[23] At the 25-percent level or better.

[24] For all but total profits and rate-of-return index.

Table 4–Multiple regression results, independent variables: General multiple regression method with B values

Independent variable	Dependent variable (related objective function)							
	Regional balance of payments	Regional balance of trade	Gross regional product	Local value added	Local aggregate wage bill	Local employment	Regional profits	Rate-of-return index
Transportation costs per dollar of export	−0.85864	−0.96942	−0.87729	−0.56811	−0.80427	−13.14893	0.04373	−0.09409
Capital/output	−.09150	−.12863	−.09543	−.08708	−.09841	−1.64322	.02202	.20949
Capital/labor	−.13545	−.14001	−.06377	−.03257	−.04126	−.18103	−.01552	−.24916
Rate of return	−.01998	.04233	−.03852	−.02886	−.01815	.32130	.02907	.76007
Value added/output	.14121	.16334	.11710	.10089	.11917	1.55377	−.01946	−.49883
Value added/labor	1.05957	1.39069	1.39285	1.16708	1.42347	18.60540	.44830	−.92049
(Managerial labor)/(total labor)	−.69029	−.84386	−.47332	−.30920	−.52896	−9.16468	−.02700	−.62353
(Skilled labor)/(total labor)	.06881	.09011	−.11507	−.12791	−.11820	−1.91295	.01255	.08998
Imported input costs per dollar of output	.04820	.05890	.05297	.04494	.05584	.97697	−.00733	.03969
Constant	.08566	.09242	.06674	.04468	.06988	1.37693	−.01603	.02834

Table 5–Multiple regression results, independent variables: General multiple regression method with BETA values

Independent variable	Dependent variable (related objective function)							
	Regional balance of payments	Regional balance of trade	Gross regional product	Local value added	Local aggregate wage bill	Local employ-ment	Regional profits	Rate-of-return index
Transportation costs per dollar of export	−0.63590	−0.57268	−0.67089	−0.56201	−0.60501	−0.58490	0.22649	−0.05261
Capital/output	−.22625	−.25372	−.24365	−.28764	−.24718	−.24405	.38083	.39108
Capital/labor	−.08999	−.07420	−.04374	−.02891	−.02784	−.00722	−.07212	−.12497
Rate of return	−.03848	.06501	−.07658	−.07423	−.03550	.03716	.39146	1.10478
Value added/output	.20043	.18492	.17162	.19128	.17181	.13246	−.19323	−.53451
Value added/labor	.15499	.16227	.21038	.22804	.21150	.16347	.45865	−.10165
(Managerial labor)/ (total labor)	−.55184	−.53811	−.39072	−.33019	−.42952	−.44006	−.15096	−.37633
(Skilled labor)/ (total labor)	.05313	.05550	−.09174	−.13193	−.09270	−.08872	.06775	.04722
Imported input costs per dollar of output	.09258	.09024	.10507	.11531	.11040	.11272	−.09851	.05756

Table 6–Multiple regression results, independent variables: General multiple regression method with F values[1]

Independent variable	Dependent variable (related objective function)							
	Regional balance of payments	Regional balance of trade	Gross regional product	Local value added	Local aggregate wage bill	Local employment	Regional profits	Rate-of-return index
Transportation costs per dollar of export	142.513	101.172	100.507	47.588	68.324	64.566	5.229	0.689
Capital/output	9.869	10.864	7.253	6.819	6.239	6.150	8.088	20.825
Capital/labor	3.007	1.789	.450	.133	.152	.010	.559	4.096
Rate of return	.330	.825	.829	.526	.149	.165	9.889	192.293
Value added/output	9.553	7.119	4.438	3.720	3.718	2.234	2.568	47.982
Value added/labor	7.625	7.316	8.902	7.057	7.520	4.542	19.313	2.316
(Managerial labor)/(total labor)	110.627	92.074	35.139	16.931	35.495	37.672	2.394	36.332
(Skilled labor)/(total labor)	.758	.724	1.433	1.999	1.223	1.132	.357	.423
Imported input costs per dollar of output	2.863	2.381	2.336	1.898	2.156	2.272	.937	.781

[1] Degrees of freedom: 1,91 for all independent variables (for all dependent variables). Critical F = 1.34 for 25-percent significance; F = 2.77 for 10 percent; F = 3.96 for 5 percent; and F = 6.96 for 1 percent.

Table 7—Multiple regression results, independent variables: General multiple regression method with BETA values and their rank (by absolute value and by column) for statistical significance of 25 percent or better

Independent variable	Dependent variable (related objective function)							
	Regional balance of payments	Regional balance of trade	Gross regional product	Local value added	Local aggregate wage bill	Local employ-ment	Regional profits	Rate-of-return index
Transportation costs per dollar of export	−0.63590 1*	−0.57268 1*	−0.67089 1*	−0.56201 1*	−0.60501 1*	−0.58490 1*	0.22649 4*	X
Capital/output	−.22625 3*	−.25372 3*	−.24365 3*	−.28764 3*	−.24718 3*	−.24405 3*	.38083 3*	.39108 3*
Capital/labor	−.08999 7*	−.07420 7*	X	X	X	X	X	−.12497 5*
Rate of return	X	X	X	X	X	X	.39146 2*	1.10478 1*
Value added/output	.20043 4*	.18492 4*	.17162 5*	.19128 5*	.17181 5*	.13246 5*	−.19323 5*	−.53451 2*
Value added/labor	.15499 5*	.16227 5*	.21038 4*	.22804 4*	.21150 4*	.16347 4*	.45865 1*	−.10165 6*
(Managerial labor)/(total labor)	−.55184 2*	−.53811 2*	−.39072 2*	−.33019 2*	−.42952 2*	−.44006 2*	−.15096 6*	−.37633 4*
(Skilled labor)/(total labor)	X	X	−.09174 7*	−.13193 6*	X	X	X	X
Imported input costs per dollar of output	.09258 6*	.09024 6*	.10507 6*	.11531 7*	.11040 6*	.11272 6*	X	X

*Rank out of nine or less.

Note: The X's denote independent variables *not* statistically significant at the 25-percent level or better.

ables. Having the largest absolute BETA value means that a variable has the largest "impact" on the generalized shadow price. That is, other things being equal, a unit change in the independent variable of one standard deviation causes the largest standard deviation change in the generalized shadow price. Using the BETA coefficient, one sees that the transport cost variable has the largest impact for most (six) of the objective functions.

Transport Cost

What might the above result for transport cost indicate? First, it may suggest that at least part of this variable's large effect on the generalized shadow price occurs because it is the only one of the nine independent variables which exhibits a different value between the two export rings for the same SIC manufacturing industry. But more than likely, both its high BETA rank, and high statistical significance for most objective functions, suggests a more fundamental reason for its indicated effect on the regional objective function. This would seem to validate, in a sense, industry location theorists' historical preoccupation with transport costs. The results observed here suggest that these costs perhaps may be more important to a planner than any other of the industry characteristics studied in this analysis.

What is the direction of the impact of the transport cost? For six of the seven maximization objectives in which transport cost is statistically significant, the regression coefficient is negative. In other words, in all but one objective function, this direction makes sense.[25] Thus, as transport costs increase, generalized shadow prices fall (worsen). Or, in other words, as transport costs fall, shadow prices rise. The conclusion for planners is that low transportation costs per dollar value of exports may be the most important characteristic of an industry being considered for location in the planning area, regardless of the development objective.

(Managerial Labor)/(Total Labor)

The second most important variable for both BETA value (ranked second in six of eight objective functions), and statistical significance (second-largest F value for six of eight objectives) is the (managerial labor)/(total labor) ratio, statistically significant for all eight objectives. In all objective functions, an increase in this ratio worsens the objective function value. The effect of this variable is easily interpretable in view of the relative shortage of managerial labor in the Adjusted Planning Model version of the RDAAP Model. One would expect an optimizing model to choose those industries which use relatively less of a scarce resource because of the relatively high opportunity costs of that resource.

The (clerical labor)/(total labor) variable is highly correlated (pairwise) with the managerial labor variable. To lessen the possibility of multicollinearity,

---

[25] In the maximization of regional rate-of-return index, this variable is not statistically significant. In the maximization of total regional profits, it is statistically significant, but the sign of the coefficient is positive (perhaps a surprising outcome). It seems difficult to explain or interpret this latter result. That is, one would have expected higher transport costs to decrease aggregate regional profits.

we deleted the clerical labor ratio. However, because of this high pairwise correlation, the partial regression coefficient for managerial labor also can be considered to apply loosely to clerical labor, as the effect of one cannot easily be separated from the other.[26]

The above effect was repeated for the (skilled labor)/(total labor) ratio, but it is statistically significant for only two objective functions–gross regional product, and local value added. Although skilled labor is not an obvious scarce resource in the model, the "shortage" type of explanation used above for managerial labor probably remains valid.

Capital/Output

Among the eight objectives for which capital/output is statistically significant, the ratio is the third most important variable in BETA value, and generally third or fourth largest in F value. A dichotomy emerges between the two profit-type criteria and the other six regional objectives. For these profit-type criteria–regional profits and rate-of-return index–as capital intensity relative to output rises, aggregate profits and index values rise.

The reverse is true for the other six objectives wherein the levels of the objective functions fall. Thus, while capital-intensive (relative to output) projects tend to improve the welfare of entrepreneurs, they lower the levels of all other regional (maximization) objectives which may be important to other groups. For these regional criteria, light or cottage-style industry may be preferable to more heavy industry.

This division of results highlights a common economic development dilemma: The most profitable ventures are not necessarily those which will improve most the welfare of the average resident of a rural area, or underdeveloped region. Investments in oil, mining, or petrochemical industries may represent several such examples of capital-intensive industries.

Value Added/Labor

The value added/labor ratio is statistically valid for all eight objectives. For all but the rate-of-return index objective, increasing this ratio results in an improved objective level.[27] Thus, for seven of the eight criteria, individually productive workers (in terms of value added) seem best for a region's economic health, as one would probably expect. The BETA value is generally ranked

---

[26] Another labor skill variable–(unskilled labor)/(total labor)–was also deleted from the list of independent variables. It is highly correlated pairwise with transportation costs, value added/labor, and profits/labor. Profits/labor also was deleted from the variable list because of its high pairwise correlations with value added/labor and transport costs. In any case, even were there no high pairwise correlations of these labor skill variables with other potential independent variables, not all *four* could be considered simultaneously. Labor is divided here into four types, and the four ratios sum to 1.0 for each industry. Thus, each percentage is an *exact* linear combination of the other three. Including all four could lead to multicollinearity problems.

[27] A dichotomy between results for profit-type and non-profit-type objectives is revealed here for only one of the two profit-type objectives–rate-of-return index. The coefficient value B (or BETA) is not, however, overwhelmingly significant for this objective, with an F of 2.316.

fourth or fifth, and F value ranked third to fifth within each of the area objective functions.

An interesting result is that value added/labor is the most important factor—ranked first for both BETA and F values—for the regional profits objective. Since wage/labor exhibits a fairly high pairwise correlation with value added/labor, wage/labor was deleted from the list of independent variables. However, if one considers value added/labor as an approximate surrogate for wage/labor, or average wage rate, the results suggest that individual industries with *higher* average wage rates are both more profitable and also the most important factor to entrepreneurs preferring increased regional profits. This result differs from a result to be discussed later for value added/output (and wage/output, or aggregate wage), wherein lower aggregate wage firms seem to be more profitable (for both profit-type regional objectives).

For the local value-added objective function, the value added/labor variable, while statistically significant, is only the fourth-ranked independent variable in BETA value. Thus, one sees that an individually high industry ratio for a given economic characteristic is not necessarily the most important determinant of a corresponding high aggregate regional value for that same economic characteristic. A similar result is observed for the value added/output variable to be discussed in the next subsection.

In short, this average wage variable tells the regional planner that the best industries for a rural planning region are those which have higher average wage rates. Thus, the typical "shirt factory" industry with its low average wage rate, which is often attracted to rural and southern areas of the United States, does not seem to be the preferred type of development for most regional objectives, including maximizing aggregate regional profits. The result for regional profits is surprising—why, then, would firms paying low wages locate in rural areas if they could not maximize profits? One explanation is that, as was said before for *all* results for the partial coefficients, the results hold precisely only when the levels for the other independent variables remain unchanged. Perhaps, in reality, such firms have characteristics which more than offset the partial effect of their low average wage. Also, the objective function measures *aggregate* profits for the region, not individual industry profits or rates of return. Perhaps what is true for individual industries may not necessarily be true for industries as a group. The fact that the relationship between value added/labor and the rate-of-return index for the region is negative, provides some support for the latter explanation.

Value Added/Output[28]

The value added/output variable generally has BETA ranks and F values similar to those for the value added/labor ratio. The value added/output ratio,

[28] The labor/output variable has been deleted from the list of explanatory variables because it is an *exact* ratio of two included variables—value added/output divided by value added/labor. Similarly, the ratio of capital/output to capital/labor would seem to equal labor/output. This is not so, however, because labor from the "total" account was used for labor/output and value added/labor, and labor from the "current" account for capital/labor (see table 1). For this same reason, the four included variables are not functionally interrelated. That is: $(VA/O)/(VA/L_T) \neq (K/O)/(K/L_C)$.

too, is significant for all eight objective functions. For all non-profit-type objectives, higher levels of value added/output lead to increased (improved) generalized shadow prices. However, for the two profit-type regional objectives—profits and rate-of-return index—a higher ratio lowers both of these objectives. One way to interpret these diametrically opposite results is to recognize that the value added/output variable is extremely highly correlated pairwise with wage/output. (Wage/output was deleted from the list of independent variables because of this high correlation.) Thus, value added/output can be considered an approximate surrogate for wage/output.

All things being equal, industry profit maximizers would rather pay lower than higher aggregate wages. However, lower value added per unit of output, while implying lower aggregate wages, also likely implies lower profits per unit of output. Nonetheless, the lower unit profits must permit a larger gross industry output to be produced for the region—although not a larger gross regional product—which more than offsets the lower individual profit/output ratios for each industry.

In short, the results for the profit-type objectives contrast with those for the other six regional objectives which emphasize higher value added/output ratios and, thus, higher aggregate wage industries. Thus, profit-maximizing criteria may lead to policy recommendations which will lower the levels of all other regional (maximization) criteria.

Imported Input Cost

The imported input cost is statistically significant for all objectives other than the two profit-type objectives.[29] In these six objectives, increasing the share of imported inputs relative to output improves the objective function values. The relative impact (BETA) of this variable—imported inputs/output—is generally fairly low among the significant variables, for all six objectives. Similarly, the F values are generally the lowest among those variables which are significant for many different regional objectives.

This finding for the imported input cost perhaps can be interpreted by considering that this result tends to support the advice of Hirschman (*1*). He suggests that, for an underdeveloped country, a planner should pursue a policy of attracting industry in which merely the "finishing touches" are put on the disassembled imported goods before they are re-exported. This implies that a very large percentage of the total value exported should consist of goods (inputs) which were previously imported. Viewed in another sense, Hirschman favors unbalanced rather than balanced growth. The latter type of growth often indicates an industry "complex" in which the total goods imported by the entire complex relative to its output are a relatively small share. Or, one can view these results in terms of relative specialization of the "world"[30] economy. That is, the results prescribe more specialization by the region as a result of the increased proportion of imported inputs.

The import cost ratio used in this analysis does not embody fully the entire level of imported inputs. The RDAAP Model can import not only goods unable to be produced by the model (the imported input cost ratio used), but

[29] This again represents a dichotomy between the results of profit-type regional objectives versus all others.

[30] The BMW Region compared with the rest of the world.

also can import goods whose production by the model is possible but not chosen in a particular optimal solution. Nonetheless, it seems probable that for most industries, the sum of the imported inputs not included in the present import cost ratio, would change that ratio only slightly. Thus, the import cost ratio used, albeit partially incomplete, should constitute a reasonably accurate surrogate for the cost ratio containing all imported inputs.

Other Industry Characteristics

The observations which can be made on the remaining industry economic characteristics (independent variables) are not as universal, insofar as their statistical reliability over all or most of the eight objective functions. Most of these observations can be said to be valid for only a reduced number (three or less) of objective function types. The results for (skilled labor)/(total labor), significant for only two objective functions, have been discussed earlier. The fact that these variables show statistical significance in only a few objectives by no means diminishes their importance for the planner, particularly if one of those objectives is the criterion of importance for a particular region.

The capital/labor ratio is significant statistically in only three regional objective functions—balance of payments, balance of trade, and rate-of-return index. In all three cases, as the relative (to labor) requirement of capital increases, the criterion function worsens (falls). The relative sizes of the BETA impacts are quite small among the significant variables for these three regional objectives. The results here for increased capital intensity (relative to labor) are identical to a majority of the capital-intensive (relative to output) results discussed earlier. That is, increased capital intensity, whether to output or labor, generally reduces regional criterion function values. The two profit-type regional objectives for the capital/output ratio are the exception. However, for capital/labor, the finding for the profit-type objectives is consistent with the majority of the capital-intensive results. For the one regional profit-type objective which is statistically significant for capital/labor—regional rate-of-return index, less capital intensive (relative to labor) industries are more profitable and thus perhaps more desirable for potential regional investors.

These results for the profit-type objectives perhaps can be interpreted by recognizing that as capital becomes relatively more abundant, its average profit rate of return should fall if profits do not increase in step. If, for an increased capital/labor ratio, profits do not increase as rapidly as capital, and if for increased capital/output profits increase more, the profit rate of return for an industry should fall in the former case, and rise in the latter. The regional profit-type objective functions used are not strictly regional profit rates of return as the objective functions are a regional rate-of-return *index* and an *aggregate* regional profit level. The scenario posited for the relationship between capital and profit levels, however, most likely explains the results observed.

For the two "mercantilist" regional objective functions—balance of payments and balance of trade—the typical effect is observed when capital intensity is reduced for the capital/labor ratio; that is, the objective function values rise. Thus, in general, the results here suggest that a less capital intensive (relative to labor or output), or a more "cottage-industry" style of development is the preferred development scheme for most regional objective functions.

The industry profit rate-of-return ratio is significant statistically for only two regional objectives, the two profit-type objectives. In both cases, as the rate-of-return ratio increases, the objective function value improves. For the rate-of-return index objective, this result is perhaps obvious as the individual industry profit rates-of-return are included in the index itself. Here, its BETA rank is not only first for the regional rate-of-return index objective function, but it also exhibits the highest impact among all eight objective functions, for all statistically significant independent variables in table 7.[31] Similarly, its F value is the largest in the entire table.

As indicated earlier, two or more objectives felt to be of equal importance to the planner, group of planners, and the area can lead to conflict. The most common potential conflict observed in these results is that levels of industry characteristics which lead to the most profitable group of industries sometimes lead also to reduced levels of all or most nonprofit types of regional objectives. And, levels leading to the least profitable group sometimes lead to greater levels of all or most nonprofit types of regional objectives. The industry characteristics of transport cost, capital/output, and value added/output exemplify this phenomenon (table 7).

As a specific example, higher capital/output ratios and lower value added/output ratios imply lower levels of both local wage bill and local employment, but these also imply higher levels of both aggregate regional profits and also the regional profit rate-of-return index. If the planner wishes to achieve all four of these objectives simultaneously, changes in these two variables will not permit this, so the planner must consider the explicit tradeoffs among these objectives.

MULTIPLE REGRESSION RESULTS: STEPWISE REGRESSION METHOD

The second multiple regression method used is the stepwise procedure explained earlier. The results will be given for only those observations which seem to differ from the results of the general multiple regression technique. In general, however, the results are similar for the two methods. The stepwise procedure, which is set to eliminate all statistically insignificant (above the 25-percent level) independent variables, could have been used alone instead of the general method. It was not used alone because the stepwise algorithm available does not give results based on standardized variables; hence, it gives no BETA coefficients.[32]

For gross regional product, (skilled labor)/(total labor) is significant in the general procedure only; for balance-of-payments surplus and local employment, it is significant in the stepwise procedure only. Also, the imported input cost ratio is significant in the stepwise procedure, but not in the general procedure, for both profit-type objectives. The reverse is true for the local value-added objective function. For the regional profits objective in the stepwise pro-

[31] The BETA rank for the rate-of-return ratio in the regional profits objective function is second of the six variables statistically significant for that objective. Thus, increased industry rates of return, while important for maximizing aggregate regional profits, are not the most important factor, either for relative impact (BETA) or statistical F value (second largest).

[32] It would also have been possible to take only the statistically significant variables obtained from the stepwise format and rerun them using a general multiple regression format, and obtain the standardized variable results as well.

cedure, an increased imported input cost ratio leads to a lower profits level. The same import ratio increase, however, leads to a rise in the profit rate-of-return index in the stepwise procedure, a result which corresponds to that for all non-profit type objectives in the general procedure given in table 7.

In general, for most objectives and independent variables, the stepwise and general procedures yield fairly similar results. Of these, usually the approximate size and always the sign of the partial (B) coefficients are identical in the two procedures.

For several regional objectives, certain variables which were selected in earlier steps as best[33] are removed in later steps because they fail to meet the 25-percent F criterion. The capital/labor and rate-of-return ratios are the only examples. This pattern occurs for capital/labor in the gross regional product objective, and for the rate-of-return ratio in the local employment regional criterion. It occurs also for both of these independent variables in the local wage bill objective. The value added/output variable seems to be the one which, when selected in the stepwise procedure, causes these two variables to lose their statistical significance.

To ascertain that these results were not due to any multicollinearity problems (variables with high pairwise correlations were already deleted), we regressed capital/labor against the rate-of-return and value-added/output variables. Multiple R^2 was negligible at 0.0938. The explanation for these inclusion and removal patterns may be that when a variable is removed after another variable has been added in some of these cases the "new" variable, in concert with some other variable(s), measures the "same thing" as the deleted variable, and eliminates it as a significant explanatory factor.[34] These patterns show that which variables are statistically significant can depend upon the number of independent variables being considered.

CONCLUSIONS

What information can the analysis in this bulletin convey to a regional economic planner? Although these results may apply fairly widely in many different rural, multicounty areas, one should be cautious in their use beyond the specific application given here. In other words, the *type* of analysis presented here should be stressed perhaps more than any specific observation or result. Such an observation or result may depend at least partially upon the study

---

[33]The method selects for inclusion the best variable (of those which meet or exceed the specified F level) from the nonincluded variables of the previous step.

[34]Another combination of variables, which were thought (theoretically) to possibly exhibit a high multiple R^2 was tested as a precaution for multicollinearity problems. When VA/O was regressed on L/O [= (VA/O)/(VA/L)] and K/O, the multiple R^2 equaled 0.4538, an acceptably low level. VA/O $\equiv$ (Pr/K) (K/O) + (W/L) (L/O) + IBT/O (Indir.Bus.Tax) for each industry. Thus, if Pr/K, W/L, and IBT did not vary also for each industry and were fixed values, the multiple R^2 would have equaled 100 percent. However, all three latter variables are at different levels for each industry. Further, as examined for one of the eight objectives, gross regional product, the "multicollinearity effect" statistic, generated by the regression algorithm, did not suggest any serious multicollinearity problems, with a value of 0.2791.

region itself (BMW Region) and the assumptions in the model used (RDAAP). There is no way to be certain of this without rerunning the model for another area, or using the technique on a different, but similar type of model.

What options does the use of this multiple regression technique in a linear programming context give the planner? He or she can estimate directly the regression equation(s) for the specific region and model, and insert the values for the independent variables (if estimable) for an industry "candidate" not already in the linear programming model. This operation would lead to an estimated generalized shadow price for that industry.

If comprehensive data for independent variables are not easily available, the planner can make ordinal (rather than cardinal) comparisons among those same candidate industries, using (in the partial sense discussed previously) the individual sample partial regression (B) coefficients. For example, if the planner is maximizing gross regional product, he or she would select an industry which, other things being equal, exhibited a lower capital/output ratio than another industry because the inclusion of this industry in the area would, per unit of output, cause a greater expansion in gross regional product. If standardized variable results are used (BETA regression coefficients), more meaningful comparisons can be made among both dependent and independent variables.

It is this researcher's contention that comparison analysis will likely be more meaningful than analysis which attempts to estimate the generalized shadow prices precisely. That is, comparison analysis represents more an ordinal sense and exact estimation, more a cardinal sense. It is doubtful in any case whether the precision in estimation of independent variables will be accurate enough to render precise shadow price estimation. Also, the time and/or cost involved may be excessive. In general, however, most planners probably would not require such precision, and this regression method can provide them with less specific "rules of thumb." For example, they may wish to know (tables 4-7) that both the aggregate local wage bill and total local employment increase, other things being equal, by introducing industries with low transport requirements, but they may not need to know the precise sizes of those increases.

In a more theoretical sense, one can say that the linear programming economic planning model gives optimal results which describe only the shadow prices and the specific SIC industry pertaining to each such price. But the multiple regression technique used here reveals the underlying economic explanations for the behavior and results of the linear programming model near its optimal corner solution. That is, the regression analysis summarizes and interprets these results in terms of economic factors which perhaps are of more use for the regional planner than the programming output alone. Further, the linear programming output is reduced in "volume" because the focus is shifted from the number of SIC's and shadow prices in the model to more easily understood industry economic characteristics.

For a sufficiently high multiple R^2, and high statistical F values for both the whole equation and partial regression coefficients, the statistically significant characteristics perhaps may be considered causal, or at least correlative determinants of the economic behavior and results of the planning model (near the optimal corner solution). This line of reasoning, however, should not be stressed excessively, in the absence of economic theory to support it. Such conclusions are somewhat tentative or exploratory in the absence of sufficient corroborative evidence from other sources. Nonetheless, while the linear programming model would seem to illuminate just one characteristic of an industry—the SIC number or commodity type—the regression analysis expands the understanding to include nine more such factors or characteristics.

More generally, whether or not an economic-type linear programming planning model is involved, multiple regression analysis can be a useful summary of any linear programming printout, if one can devise some potential explanatory factors for the problem which can be hypothesized to have some effect on the generalized shadow prices. Such a technique might prove especially useful if the results of the programming analysis were extremely numerous and difficult to understand. In other words, multiple regression may be able to translate the linear programming output into a more meaningful form for many users of the research.

Some of the specific results—such as the recommendation of industries with low transport costs and reduced use of scarce resources (such as managerial and skilled labor)—concur with the conventional planning wisdom. But many do not. For example, it is often felt that rural areas should attempt to attract low-wage industry. The results for seven regional objectives indicate that high average wage industries may be preferable (because they are also highly labor productive). And, to the extent that such wisdom also implies that industries to be attracted should exhibit low value added and low aggregate wage per unit of output, the results here suggest that such a policy may not maximize most regional (non-profit-type) objectives.

It is also often suggested that a large infusion of capital is necessary for economic growth. While investment capital is obviously needed for regional growth, this should not imply (as it often does) that capital intensity need be great for most individual industries for an area to achieve many aggregate goals. In fact, the negative relationship between capital/output and most regional objectives tends to imply the converse. In addition, some economic development strategies for less developed regions have stressed the need for industrial complexes, with their implied low import requirements. These complexes have been suggested as being best for the planning area so that it can achieve more balanced growth and independence from other areas. The results here suggest that higher import requirements are associated with larger increases in most regional objectives.

In all three of the above regression results which tend to contradict the conventional planning wisdom, there is an implied opportunity cost (foregone increments in the current objective function level) in pursuing each of these conventional strategies. This cost may exceed any benefit obtained from implementing such a strategy, but this topic has not been explored.

APPENDIX

Manufacturing export commodities, related SIC number(s), actual generalized shadow price, estimated generalized shadow price, and industry economic characteristics: for the objective function of maximizing gross regional product: general multiple regression method

Export name	Related SIC number	Actual generalized shadow price[1]	Estimated generalized shadow price	Transportation costs per export dollar[1]	Capital/output	Capital/labor	Rate of return	Value added/output	Value added/labor	(Managerial labor)/(total labor)	(Skilled labor)/(total labor)	Imported input costs per output dollar
EI 2011	2011	−0.00761	0.02833	0.00761	0.13728	0.02827	0.31104	0.16093	0.03222	0.129	0.092	0.05811
2015	2015	−.01672	.01797	.01672	.17343	.03567	.24142	.15871	.03151	.128	.097	.06224
2021	2021	−.01393	−.05133	.01393	.26256	.04383	.05545	.06031	.00961	.175	.149	.06924
2022	2022, 2025	−.01022	−.02884	.01022	.34180	.04390	.12016	.15830	.01934	.184	.117	.16077
2023	2023	−.02249	−.02184	.02249	.20497	.03319	.22306	.17458	.02719	.186	.110	.07181
2024	2024	−.00182	−.03778	.00812	.53128	.06056	.09045	.20225	.02037	.184	.117	.24480
2026	2026	−.01812	.08507	.01218	.31026	.04842	.46019	.53607	.08047	.186	.109	.09744
2037	2037	−.02851	−.00297	.02851	.18881	.31026	.36396	.25799	.03035	.135	.128	.25045
2040	2041, 2043, 2045	−.07129	−.06693	.07129	.22594	.05682	.19766	.17051	.04044	.218	.189	.19472
2042	2042	−.03656	−.03857	.03656	.21091	.05520	.18602	.15125	.03771	.219	.185	.23192
2044	2044	−.07403	−.08987	.07403	.06893	.02160	.37212	.09648	.02954	.226	.167	.08250
2046	2046	−.07038	−.06020	.07038	.66140	.10717	.15080	.38504	.05586	.207	.218	.20258
2092	2092	−.08003	−.10922	.08003	.37407	.30541	.02275	.04391	.02740	.141	.275	.10156
2250	225 (excl. 2256)	.04740	−.00241	.08714	.36894	.02092	.19898	.39479	.02189	.061	.123	.42717
2256	2256	.06204	.02977	.06300	.36894	.05287	.12262	.24682	.03345	.059	.152	.59521
2320	2322, 2328	.10345	.07232	.00171	.07262	.00459	.93142	.35721	.02245	.074	.045	.32631
2421	2421	−.09769	−.07185	.10609	.61318	.03965	.17921	.36547	.02264	.092	.165	.20059
2426	2426	−.00008	.01621	.00008	.61318	.02839	.17921	.36547	.01640	.089	.151	.20059
2431	2431	.00915	.03740	.01218	.19576	.01165	.51813	.32828	.02740	.094	.146	.10231
2432	2432	−.02914	.00352	.04547	.33032	.02733	.29296	.31645	.02543	.093	.155	.20004
2440	244	−.07229	−.02169	.07229	.39680	.02121	.06295	.33650	.01759	.094	.149	.27112

Continued –

Manufacturing export commodities, related SIC number(s), actual generalized shadow price, estimated generalized shadow price, and industry economic characteristics: for the objective function of maximizing gross regional product: general multiple regression method—continued

Export name	Related SIC number	Actual generalized shadow price[1]	Estimated generalized shadow price	Transportation costs per export dollar[1]	Capital/ output	Capital/ labor	Rate of return	Value added/ output	Value added/ labor	(Managerial labor)/ (total labor)	(Skilled labor)/ (total labor)	Imported input costs per output dollar
EI 2511	2511, 2512	-0.00453	0.03630	0.00453	0.29863	0.01654	0.20229	0.37530	0.02041	0.107	0.217	0.26648
2514	2514	-.03352	.02258	.03352	.35059	.02885	.14881	.32820	.02630	.106	.222	.50838
2531	2522, 2531, 2599	.16620	.06539	.01178	.33396	.02656	.23398	.51184	.03956	.106	.224	.23548
2540	254	.09958	.05566	.00997	.29015	.02310	.22237	.42287	.03286	.106	.221	.29648
2650	265	.02116	.04261	.02595	.34880	.03450	.32210	.39493	.03767	.128	.161	.50100
2821	2821	-.01504	-.04581	.01504	.73573	.10704	.18857	.33661	.04362	.252	.252	.62325
2841	2841	-.02746	-.02086	.02746	.54080	.12936	.30368	.32302	.06739	.245	.265	.43056
2842	2842	-.01630	-.01905	.01630	.39431	.07074	.44069	.33707	.05604	.267	.222	.41563
2851	2851	-.01886	-.02321	.01886	.35166	.05231	.40417	.35904	.05037	.294	.132	.40019
2872	2872	-.00008	-.03840	.00008	.42315	.05858	.28163	.24431	.03182	.272	.210	.56403
2879	2879	-.05005	-.06753	.05005	.50494	.09647	.23906	.25249	.04276	.250	.260	.60997
2951	2951	-.04259	-.07234	.04259	1.03509	.14349	.02586	.30026	.03613	.162	.254	.52153
2952	2952	-.17199	-.13908	.17199	.37015	.04088	.08392	.24225	.02561	.184	.181	.56844
3111	3111	.04710	.05230	.00257	.30120	.03326	.15010	.25976	.02766	.088	.127	.22241
3141	3141	.22362	.09955	.00189	.17171	.00954	.45554	.46966	.02583	.056	.065	.22665
3241	3241	-.40894	-.43391	.45566	1.45146	.16129	.11169	.50164	.04766	.137	.271	.43593
3250	325	-.19648	-.14593	.21143	.77705	.04543	.23193	.53715	.03001	.101	.151	.36396
3411	3411	.00770	.02144	.01257	.51303	.07015	.15354	.33465	.04259	.147	.292	.53666
3421	3421	-.00203	.02836	.00203	.29667	.04057	.34475	.37699	.03528	.155	.264	.34053
3423	3423, 3425	-.00709	.03585	.00709	.37769	.02698	.36366	.50569	.03512	.156	.263	.43549
3429	3429	-.00272	.02532	.00272	.43085	.03375	.26425	.42169	.03193	.154	.271	.46504
3431	3431, 3432	-.01364	.00163	.01364	.53091	.04511	.13434	.35538	.02890	.153	.273	.51985

Continued –

Manufacturing export commodities, related SIC number(s), actual generalized shadow price, estimated generalized shadow price, and industry economic characteristics: for the objective function of maximizing gross regional product: general multiple regression method—continued

Export name	Related SIC number	Actual generalized shadow price[1]	Estimated generalized shadow price	Transportation costs per export dollar[1]	Capital/ output	Capital/ labor	Rate of return	Value added/ output	Value added/ labor	(Managerial labor)/ (total labor)	(Skilled labor)/ (total labor)	Imported input costs per output dollar
EI 3433	3433	−0.00975	0.01884	0.00975	0.31859	0.03382	0.20286	0.31826	0.03271	0.155	0.266	0.41850
3585	3585	.14567	.05979	.00703	.28895	.03627	.38951	.49638	.06012	.163	.325	.31706
3629	3629	−.00257	.01324	.00257	.19605	.01346	.41163	.40529	.02740	.227	.180	.50423
3691	3691	−.01019	.01521	.01019	.41679	.04529	.29360	.42233	.04389	.218	.202	.51380
3694	3694	−.00257	.00743	.00257	.29751	.02248	.41111	.41960	.03098	.224	.186	.44835
3791	3791	.05179	.03808	.00603	.32265	.03857	.04063	.27651	.03183	.112	.309	.43993
3821	3821	−.00195	−.01383	.00195	.28570	.02034	.33245	.44428	.03092	.273	.220	.41438
3822	3822	−.00054	−.01570	.00054	.33115	.02004	.32499	.50375	.02981	.273	.220	.32954
3840	3841, 3842, 3843	−.00187	−.01446	.00187	.25263	.01740	.38309	.45168	.03050	.274	.217	.37201
3941	3941	−.00091	.03462	.00091	.24041	.01538	.37627	.39266	.02471	.149	.173	.32329
3949	3949	−.00082	.04713	.00082	.17237	.01240	.59604	.44350	.03147	.147	.184	.30765
3950	395	.00878	.06217	.00177	.34792	.03150	.29728	.45054	.03292	.122	.155	.38918
3981	3981	−.00779	−.00520	.00779	.71899	.06262	.10540	.34088	.02789	.141	.208	.44300
EO 2011	2011	−.01487	.02291	.01487	.13728	.02827	.31104	.16093	.03222	.129	.092	.05811
2015	2015	−.03229	.00431	.03229	.17343	.03567	.24142	.15871	.03151	.128	.097	.06224
2021	2021	−.02691	−.06272	.02691	.26256	.04383	.05545	.06031	.00961	.175	.149	.06924
2022	2022, 2025	−.01973	−.03718	.01973	.34180	.04390	.12016	.15830	.01934	.184	.117	.16077
2023	2023	−.01455	−.03856	.04155	.20497	.03319	.22306	.17458	.02719	.186	.110	.07181
2037	2037	−.05211	−.02368	.05211	.18881	.31026	.36396	.25799	.03035	.135	.128	.25045
2040	2041, 2043, 2045	−.12714	−.11592	.12714	.22594	.05682	.19766	.17051	.04044	.218	.189	.19472
2044	2044	−.13203	−.14076	.13203	.06893	.02160	.37212	.09648	.02954	.226	.167	.08250

Continued –

Manufacturing export commodities, related SIC number(s), actual generalized shadow price, estimated generalized shadow price, and industry economic characteristics: for the objective function of maximizing gross regional product: general multiple regression method—continued

Export name	Related SIC number	Actual generalized shadow price[1]	Estimated generalized shadow price	Transportation costs per export dollar[1]	Capital/ output	Capital/ labor	Rate of return	Value added/ output	Value added/ labor	(Managerial labor)/ (total labor)	(Skilled labor)/ (total labor)	Imported input costs per output dollar
EO 2046	2046	−0.12251	−0.10857	0.12551	0.66140	0.10717	0.15080	0.38504	0.05586	0.207	0.218	0.20258
2250	225 (excl. 2256)	−.02263	−.05903	.15717	.36894	.02092	.19898	.39479	.02189	.061	.123	.42717
2256	2256	.01101	−.01499	.11403	.36894	.05287	.12262	.24682	.03345	.059	.152	.59521
2320	2322, 2328	.10203	.07107	.00313	.07262	.00459	.93142	.35721	.02245	.074	.045	.32631
2421	2421	−.17909	−.14327	.18749	.61318	.03965	.17921	.36547	.02264	.092	.165	.20059
2426	2426	−.00013	.01616	.00013	.61318	.02839	.17921	.36547	.01640	.089	.151	.20059
2431	2431	.00000	.02906	.02133	.19576	.01665	.51813	.32828	.02740	.094	.146	.10231
2432	2432	.06402	−.02708	.08035	.33032	.02733	.29296	.31645	.02543	.093	.155	.20004
2511	2511, 2512	−.00917	.03223	.00917	.29863	.01654	.20229	.37530	.02041	.107	.217	.26648
2514	2514	−.06789	−.00757	.06789	.35059	.02885	.14881	.32820	.02630	.106	.222	.50838
2531	2522, 2531, 2599	.15413	.05481	.02385	.33396	.02656	.23398	.51184	.03956	.106	.224	.23548
2540	254	.08937	.04670	.02018	.29015	.02310	.22237	.42287	.03286	.106	.221	.29648
2821	2821	−.02650	−.05589	.02650	.73573	.10704	.18857	.33661	.04362	.252	.252	.62325
2841	2841	−.04925	−.03974	.04925	.54080	.12936	.30368	.32302	.06739	.245	.265	.43506
2842	2842	−.02924	−.03041	.02924	.39431	.07074	.44069	.33707	.05604	.267	.222	.41563
2851	2851	−.03457	−.03700	.03457	.35166	.05231	.40417	.35904	.05037	.294	.132	.40019
2872	2872	−.00014	−.03845	.00014	.42315	.05858	.28163	.24431	.03182	.272	.210	.56403
2879	2879	−.09080	−.10328	.09080	.50494	.09647	.23906	.25249	.04276	.250	.260	.60997
3111	3111	.04500	.05046	.00467	.30120	.03326	.15010	.25976	.02766	.088	.127	.22241
3141	3141	.22211	.09823	.00338	.17171	.00954	.45554	.46966	.02583	.056	.065	.22665
3421	3421	−.00377	.02684	.00377	.29667	.04057	.34457	.37699	.03528	.155	.264	.34053

Continued –

Manufacturing export commodities, related SIC number(s), actual generalized shadow price, estimated generalized shadow price, and industry economic characteristics: for the objective function of maximizing gross regional product: general multiple regression method–continued

Export name	Related SIC number	Actual generalized shadow price[1]	Estimated generalized shadow price	Transportation costs per export dollar[1]	Capital/ output	Capital/ labor	Rate of return	Value added/ output	Value added/ labor	(Managerial labor)/ (total labor)	(Skilled) labor)/ (total labor)	Imported input costs per output dollar
EO3423	3423, 3425	−0.01295	0.02518	0.01295	0.37769	0.02698	0.36366	0.50569	0.03512	0.156	0.263	0.43549
3429	3429	−.00499	.02333	.00499	.43085	.03375	.26425	.42169	.03193	.154	.271	.46504
3431	3431, 3432	−.02494	−.00828	.02494	.53091	.04511	.13434	.35538	.02890	.153	.273	.51985
3433	3433	−.01781	.01177	.01781	.31859	.03382	.20286	.31826	.03271	.155	.266	.41850
3585	3585	.14022	.05501	.01248	.28895	.03627	.38951	.49638	.06012	.163	.325	.31706
3629	3629	−.00469	.01138	.00469	.19605	.01346	.41163	.40529	.02740	.227	.180	.50423
3691	3691	−.01860	.00783	.01860	.41679	.04529	.29360	.42233	.04389	.218	.202	.51380
3694	3694	−.00469	.00557	.00469	.29751	.02248	.41111	.41960	.03098	.224	.186	.44835
3791	3791	.04679	.03370	.01103	.32265	.03857	.04063	.27651	.03183	.112	.309	.43993
3821	3821	−.00350	−.01519	.00330	.28570	.02034	.33245	.44428	.03092	.273	.220	.41438
3822	3822	−.00098	−.01608	.00098	.33115	.02004	.32499	.50375	.02981	.273	.220	.32954
3840	3841, 3842, 3843	−.00338	−.01578	.00338	.25263	.01740	.38309	.45168	.03050	.274	.217	.37201
3941	3941	−.00183	.03381	.00183	.24041	.01538	.37627	.39266	.02471	.149	.173	.32329
3949	3949	−.00169	.04637	.00169	.17237	.01240	.59604	.44350	.03147	.147	.184	.30765
3950	395	.00705	.06065	.00350	.34792	.03150	.29728	.45054	.03292	.122	.155	.38918
3981	3981	−.01495	−.01148	.01495	.71899	.06262	.10540	.34088	.02789	.141	.208	.44300

[1] For the gross regional product, regional balance-of-payments surplus, and regional balance-of-trade surplus objective functions, the actual generalized shadow price exactly equals the transportation costs for exports for the manufacturing industries not in the optimal solution, but does not for those industries in the optimal solution. Thus, the least "expensive" (in terms of the objective function) method by which the model can "produce" one unit of export for a nonoptimal export industry, is to import one unit for a dollar, and then export that same unit and lose the transportation charge on the dollar of export. For no other area objective functions (of the eight considered here), does this same correspondence hold. As transportation charge is an important factor for most of the regional objective functions in terms of F and BETA values, this correspondence in just these three objectives cannot be considered the crucial factor–whatever that factor might be–which establishes the importance of transportation costs in this analysis.

REFERENCES

[1] Hirschman, Albert O. *The Strategy of Economic Development.* New Haven: Yale University Press, 1958.

[2] Johnston, J. *Econometric Methods.* New York: McGraw-Hill, Inc., 1963.

[3] Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent. *Statistical Package For The Social Sciences 2nd ed.*, New York: McGraw-Hill, Inc., 1975.

[4] Spiegelman, Robert G., E. L. Baum, and L. E. Talbert. *Application of Activity Analysis to Regional Development Planning: A Case Study of Economic Planning in Rural South Central Kentucky.* U.S. Dept. Agr., Econ. Res. Serv., Tech. Bul. 1339, Mar. 1965.

[5] Williams, Daniel G. "Objective Function Tradeoff Curves In a Rural Economic Development, Activity Analysis, Linear Programming Planning Model." Presented at the Western Regional Science Association Meetings, Monterey, Calif., Feb. 21-23, 1980.

[6] ________________. "Agricultural Census Data As a Source of Linear Programming Vectors," *Agr. Econ. Res.*, Vol. 30, No. 2, Apr. 1978, pp. 34-37.

[7] ________________. "Survey of the Literature Relevant to Linear Programming, Economic Development Planning Models: Relationship to a Specific Rural Area Model (RDAAP)." Working Paper No. 7905, Econ. Dvlpt. Div., Econ. Statis. Coop. Serv., U.S. Dept. Agr., July 1979.

[8] ________________. *Economic Planning for Multicounty Rural Areas: Application of a Linear Programming Model in Northwest Arkansas.* (In process.)

[9] ________________. "Structural Details of a Linear Programming, Rural Economic Development Planning Model: Northwest Arkansas." Working Paper No. 7907, Econ. Dvlpt. Div., Econ. Statis. Coop. Serv., U.S. Dept. Agr., June 1979.

[10] ________________. "Summarizing and Interpreting Linear Programming Solutions Through Multiple Regression Analysis: Application to a Rural Economic Development Planning Model." Presented at 11th annual meeting, Mid-Continent Regional Science Association, Minneapolis, Minn., May 31-June 2, 1979.

UNITED STATES DEPARTMENT OF AGRICULTURE
WASHINGTON, D.C. 20250

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF
AGRICULTURE
AGR 101
THIRD CLASS

Economics, Statistics, and Cooperatives Service

The Economics, Statistics, and Cooperatives Service (ESCS) collects data and carries out research projects related to food and nutrition, cooperatives, natural resources, and rural development. The Economics unit of ESCS researches and analyzes production and marketing of major commodities; foreign agriculture and trade; economic use, conservation, and development of natural resources; rural population, employment, and housing trends, and economic adjustment problems; and performance of the agricultural industry. The ESCS Statistics unit collects data on crops, livestock, prices, and labor, and publishes official USDA State and national estimates through the Crop Reporting Board. The ESCS Cooperatives unit provides research and technical and educational assistance to help farmer cooperatives operate efficiently. Through its information program, ESCS provides objective and timely economic and statistical information for farmers, government policymakers, consumers, agribusiness firms, cooperatives, rural residents, and other interested citizens.